This document is downloaded from the
Digital Open Access Repository of VTT

Title	Reliability of an in-vehicle warning system for railway level crossings a user-oriented analysis
Author(s)	Öörni, Risto
Citation	IET Intelligent Transport Systems vol. 8(2014):1, pp. 9 - 20
Date	2014
URL	http://dx.doi.org/10.1049/iet-its.2012.0129
Rights	Copyright 2014 The Institution of Engineering and Technology. This is a post-print version of the published article and it may be downloaded for personal use only

VTT http://www.vtt.fi P.O. box 1000 FI-02044 VTT Finland	By using VTT Digital Open Access Repository you are bound by the following Terms & Conditions. I have read and I understand the following statement: This document is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of this document is not permitted, except duplication for research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered for sale.

Reliability of an in-vehicle warning system for railway level crossings – a user-oriented analysis

Mr. Risto Öörni
VTT Technical Research Centre of Finland
P.O. Box 1000
FI-02044 VTT
Finland

e-mail: risto.oorni@vtt.fi
GSM: +358 40 7242614

ABSTRACT

The study analyses the reliability of an in-vehicle warning system developed in Finland during 2008-2010. The system is based on the positioning of trains using GPS, calculation of the states of level crossings on a server, and in-vehicle equipment which retrieves information about the states of level crossings from the server. Information about the reliability of the system is very relevant for accurate estimation of the impacts of the system and for estimating the potential for improvements to it. The study starts with a description of the system under analysis, continues by defining the concept of reliability, and provides an estimate for the reliability of the system from the user point of view. To achieve this objective, the study defines the relevant concepts and describes a methodology for analysis of the reliability of the system. The main input to the analysis of reliability includes a brief overview of existing concepts related to the reliability of in-vehicle ITS systems and empirical data obtained in a field test carried out in southern Finland. The analysis shows that the expected functionality has been achieved, but the reliability level of the pilot system needs improvement, especially reduction in the share of missed alarms.

1. INTRODUCTION

Background

The safety of level crossings is a major safety issue for railways and an existing traffic safety problem for road users. According to statistics published by the European Railway Agency (ERA), at least 123 000 level crossings exist in the EU [1]. Most of them (71%) are passive level crossings without any active warning or protection devices such as lights, bells or gates. Roughly 45% of level crossing accidents in the EU occur at passive level crossings, and 65% of road users involved in accidents are drivers or occupants of passenger cars or heavy vehicles [1]. In 2010 there were 359 level crossing accident fatalities in the EU [1]. This represents 29% of fatalities in railway accidents but only about 1.2% of all road accident fatalities [1]. Most of the direct causes are related to the behaviour of road users such as distraction [1]. Other causes of accidents were also identified, such as weather conditions or the condition of the driver (e.g. intoxication with alcohol or drugs).

An evaluation of accident data on 256 level crossing accidents was carried out as part of the SELCAT (Safer European Level Crossing Appraisal and Technology) project. About 91% of level crossing accidents in the EU were found to be caused by human failure [2], and over 80% were found to have been caused by the driver of the road vehicle not respecting the traffic rules [2].

Several countries still have unprotected level crossings on their railway networks. For example, the Finnish railway network has over 3800 level crossings in total, over 3000 of which are unprotected and have no bells, lights or gates [3]. On average, close to 10 people die in level crossing accidents every year in Finland and over 17 are injured [3, 4]. According to the Finnish Accident Investigation Centre, most accidents at passive level crossings in Finland are caused by the vehicle driver misjudging the situation but not by intentional risk-taking [3]. In other words, most of the accidents are caused by unintentional driver error. This suggests that at least some of these accidents could be avoided if the awareness and alertness of road users were increased.

The number of people injured or killed in road accidents depends on three factors: exposure, accident rate and injury severity [5]. The in-vehicle warning system for level crossings aims to reduce the accident rate by increasing the alertness and situation-awareness of the road user in situations where he/she is approaching a passive level crossing. Upon receiving a warning of an approaching train, the vehicle driver would be more likely to exercise greater caution or to stop the vehicle before crossing the tracks. Among other things, the impact of the system on the behaviour of road users depends on the reliability of the system.

A comparison of the in-vehicle warning system analysed in this paper with other solutions providing a warning of an approaching train is presented in Table 1. The table compares available systems and those under development in terms of cost and reliability and provides examples of these systems. In-vehicle warning systems for railway level crossings have been under development in recent years at least in Finland, the US and Australia [6-10]. While different technical solutions can be used to warn road users of approaching trains, the in-vehicle warning system analysed in this study has certain advantages. First, it requires no installation of equipment either on the tracks or at level crossings, which is especially useful for railway systems with large numbers of passive level crossings but few users where such solutions may be economically unfeasible. Second, it uses radio technologies and communication protocols which are standardised and widely supported, making integration with existing end-user terminals such as mobile devices or in-vehicle units relatively easy. Implementing the in-vehicle part of the system on the same platform with other ITS applications will also reduce costs.

Table 1. Systems that warn road users of approaching trains.

System	Reliability	Costs	Examples	Notes
Level crossing equipment with visual warning, connected with cables to railside systems	Usually high Finland [11]: MTBF >= ~2.2x10^5 h MTTR < 12 h SIL 3	Moderate installation costs per level crossing, low or moderate operating and maintenance costs (USA: 40 000 $/level crossing [12], higher in Europe)		Provides warning to all road users, may be equipped with bells, powered from mains electricity, usually activated by a track circuit or other railside system
Level crossing system with visual warning and a physical barrier	High Finland [11]: MTBF >=	Moderate to high installation costs per level crossing, moderate operating and maintenance costs		Provides warning to all road users, usually activated by a track circuit or other railside system

(such as boom gates)	~2.2x10^5 MTTR < 12 h SIL 3	(USA: 250 000$-350 000$/level crossing [12])		
Low-cost level crossing equipment with visual warning	Not evaluated for all systems; reliability targets for systems exist from SIL1 to SIL3 [13]	Low to moderate installation costs per level crossing Target cost [13]: 52 600-105 300$/level crossing has been mentioned for Australia. Target costs from 15 000$ to 67 000$ have been mentioned for recently developed systems or systems in the prototype stage [13].	MICRO, SAL0, HRI-2000, HiLux, O'Conner, EVA 3000, ISIS-EK (see [13] for comparison)	Low-cost solutions still under development. May be powered from mains electricity or use alternative solutions. Various activation methods. Provides warning to all road users. Legal issues especially related to liability are an obstacle for deployment, and installation of equipment at every level crossing may be costly for networks with a large number of level crossings.
In-vehicle systems based on wireless V2V communication	Not evaluated, reliability will be heavily dependent on performance of the V2V data link	No roadside equipment needed, costs of train units and in-vehicle units are dependent on implementation of the system.	Most likely to be based on IEEE802.11p	Both communicating parties need interoperable hardware and software. V2V radio technologies such as IEEE802.11p have been standardised but no large-scale integration to end-user terminals or vehicles has yet taken place.
In-vehicle warning systems implemented with short-range radio technologies (V2I)	Not evaluated / unsatisfactory [14]	Cost of level crossing equipment can be expected to be similar to low-cost warning systems. Cost of in-vehicle system varies between systems.	[7, 14], [9, 10], [8]	Requires installation of hardware at level crossing. Proprietary radio technologies used by prototype systems may be unavailable or hard to integrate with end-user terminals or in-vehicle systems.
In-vehicle warning system utilizing mobile network for data transmission (V2I)	Under analysis (this study)	Cost of in-vehicle units depends on possibilities to share the in-vehicle platform with other ITS applications. Investment required for data collection and service provision estimated at about 2 M€ for Finland [15]. Operating and maintenance costs estimated at about 0.6 M€/year [15].	Junavaro [15]	Requires coverage of public land mobile network, equipment on train and a suitable in-vehicle system but no equipment at level crossings. Advantageous in cases with a large number of passive level crossings with few users. Relatively easy integration with existing end-user terminals such as in-vehicle devices or mobile terminals.

The functionality and operating principle of the system analysed in this study was described on a general level by Öörni and Virtanen [6] after a small-scale proof-of-concept test carried out in 2006. A description of the system is provided in this paper and elsewhere [15].

Objectives

The objective of the study is to analyse the reliability of an in-vehicle warning system for railway level crossings from the user point of view. Our system was developed and tested in Finland during 2008-2011. This included collecting data during a field test involving observations on the output of the system, analysing the results, and providing a quantitative estimate for reliability of the system. The reliability was analysed from the user point of view for two main reasons: First, it gives an overall picture of the current reliability level of the system and provides guidance on the factors contributing to it. Second, the reliability experienced by end-users largely determines the potential of the system to improve their safety. Additionally, having a definition for reliability from earlier research helped define the scope of the study. Overall, the analysis of reliability is an integral part of the development process providing useful information to developers for improving the system, and information on reliability is needed before testing the system with end-users to study its impacts.

Structure of the paper

The following section provides a system description and describes requirements for system reliability. Section 3 describes the methods used to analyse system reliability. Section 4 documents the way the methods have been applied in this case and the tools used for data collection. The test results are presented in section 5 and concluding remarks and recommendations for further research in section 6.

2. SYSTEM DESCRIPTION

Functionality of the implemented system

The basic functionality of the system is to provide a warning of an approaching train to the driver of a car nearing a level crossing. The in-vehicle system has the coordinates of all level crossings in tabular format. As the vehicle moves along the road network, the in-vehicle system continuously calculates distances between the vehicle and the nearest level crossings using GPS positioning and the level crossing coordinates to detect situations in which the car is approaching a level crossing. When an approaching situation is detected, the in-vehicle device sends a query to the server with the number of the level crossing. The server receives continual real-time information on the positions of trains on the rail network and calculates the status of level crossings accordingly. When a query is received with the number of a particular level crossing, the server responds with the current status of the level crossing as: "No information", "Coming", "Alarm" or "Passed". If the in-vehicle system that has sent the query receives the response "Alarm", it warns the driver both visually and audibly. The basic architecture of the system is illustrated in Figure 1.

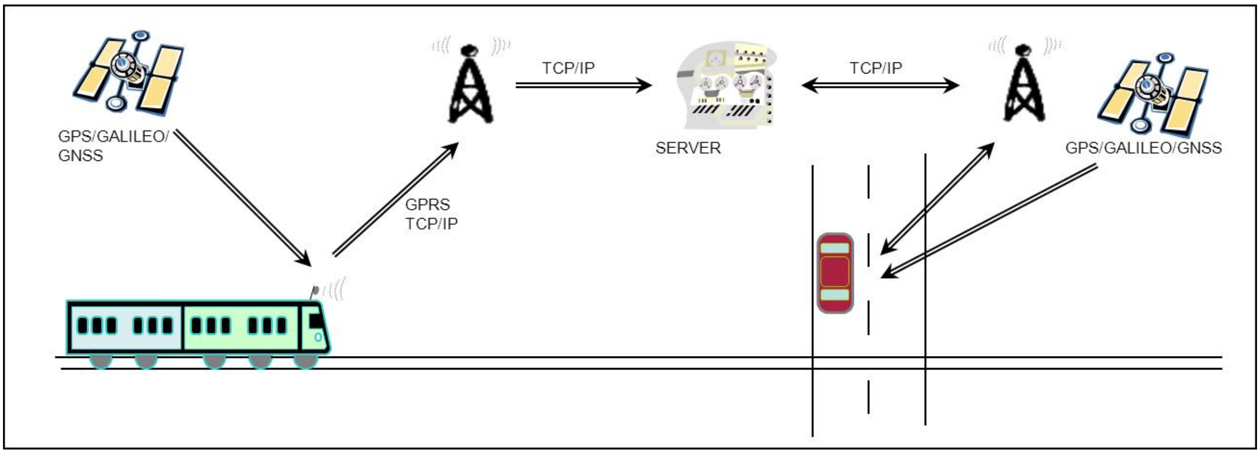

Figure 1. In-vehicle warning system for railway level crossings – simplified technical architecture [6].

The warning functionality provided by the system is based on both the distance between the vehicle and the level crossing and the direction in which the vehicle is travelling. If the distance between the vehicle and level crossing is less than 1500 metres but more than 200 metres and the vehicle is travelling towards the level crossing, the in-vehicle device warns the driver. If the distance between the level crossing and vehicle is less than 200 metres, the direction of travel has no effect and the system issues a constant audible warning with a continuous and visual indication if the status of the level crossing is “alarm”.

Functional requirements

The high-level operating requirements for the system were divided into two categories: basic functionality and other requirements. Three main requirements were identified for basic functionality:

- The driver receiving a warning must have enough time for braking
- Provision of a clear warning the meaning of which is easy to understand
- Provision of relevant information: the right information in the right place at the right time to the right users

Requirements for system reliability

A clear description of the acceptable level of reliability is needed to answer the question of whether the system is sufficiently reliable for its purpose. Because no directly applicable guidelines or standards were available, the level of reliability considered acceptable for the system was determined on the basis of a literature study and comparison with systems providing comparable functionalities within the same context.

The system provides a warning of an approaching train to the driver, but it neither intervenes in the driving task nor provides traffic control functionalities. It supports the driver by providing a warning in a potentially unsafe situation, but the driver has no obligation to comply with the information or even to use the system. The system will support the driver especially at passive level crossings, and is not intended to replace existing solutions where they exist. The system can be considered to be a driver assistance application providing a warning to the driver in potentially unsafe situations.

According to a proposal for statement of principles on the design of high-priority warning signals for in-vehicle ITS, published by the ITS Informal Group of UNECE (United Nations Economic Commission for Europe), the rate of missed alarms and false positives should be kept to a minimum and ADAS (advanced driving assistance systems) must not compromise safety [16]. On the other hand, perfect system performance is not seen in the document as a realistic objective for many systems.

Simulator studies with an in-vehicle forward collision warning system have shown that even imperfect warning systems may be useful and lead to safer behaviour [17]. The results of another study on forward collision warning systems suggest that a system with a 60% false alarm rate (false alarms / all alarms) or more was not effective [18]. The authors considered the possibility that the effect was related to system distrust caused by false alarms.

Guidelines on system performance have also been proposed in the US for integrated vehicle-based safety systems (IVBSS) providing warnings to the driver [19]. The guidelines are based on the functional requirements presented in the reports of the US IVBSS project and ISO standards efforts. They state that an IVBSS such as blind spot monitoring and warning or lane departure warning should be able to provide a warning 90% of the travel time.

The effects of the reliability of warnings provided by level crossing warning systems to drivers have been analysed in a study by Gil, Multer and Yeh [20]. They concluded that when drivers perceive a warning system to be less reliable, they are less willing to comply with the warning — even though they found it impossible to define empirically the precise warning reliability needed to achieve a given level of compliance. However, this study focused on existing level crossing equipment instead of ADAS.

On the basis of the studies above, the acceptable level of reliability in terms of true positive rate was considered to be 90%. In terms of false positive rate, only a limited amount of information was available. The acceptable reliability for the system in terms of the share of false alarms of all alarms was assumed to be 60% or less with high certainty.

Technical architecture

The technical architecture of the system used in the field test is illustrated in Figure 2 to provide an overview of the way the system has been realised and the potential factors contributing to its reliability. The level crossing server calculates the status of the level crossings based on real-time positioning information received from trains equipped with MC40, CL341 and A1 Trax devices.

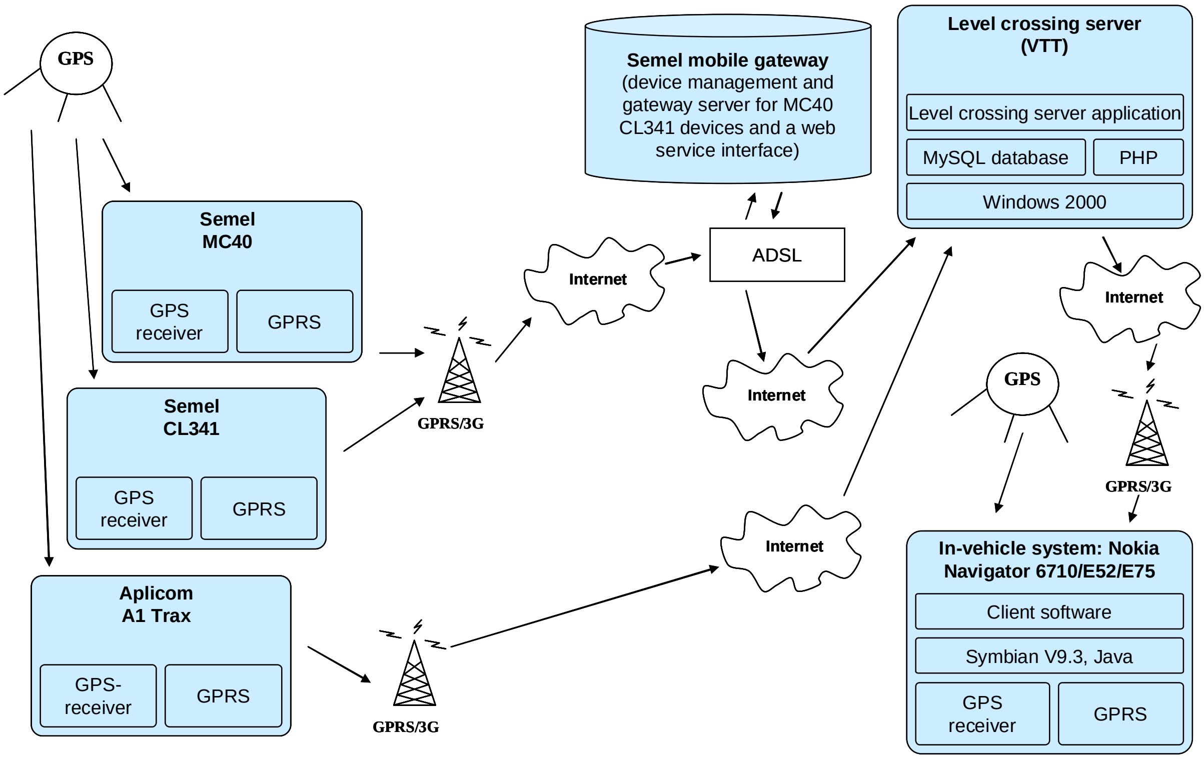

Figure 2. Technical architecture of the in-vehicle warning system for railway level crossings used in the field test.

The in-vehicle system used in the field test was a commercially available navigator phone Nokia 6710 Navigator. The phone has a built-in GPS receiver, support for AGPS (assisted GPS) service and capabilities for GPRS (general packet radio service) and UMTS (universal mobile telecommunication system) data transmission.

Three types of positioning devices installed in trains were used in the field test: A1 Trax, CL341 and MC40. A1 Trax is a tracking and security unit designed for in-vehicle use in road transport. CL341 and MC40 are in-vehicle telematics units designed for road traffic applications. A1 Trax devices communicated with the level crossing server using TCP (transmission control protocol) as a transport protocol for positioning messages. MC40 and CL341 communicated with the mobile gateway using a proprietary transport protocol. The positioning data transmitted by CL341 and MC40 devices was received by a dedicated mobile gateway server and then forwarded to the level crossing server.

3. METHODS FOR ANALYSING RELIABILITY

Overview

This section describes the methods used to assess the reliability of the system. It starts with definition of reliability and reliability measures and continues with descriptions of the methods used to analyse the reliability of the system. Data collection and modelling of the system to be analysed are described in section 4.

Definition of reliability and reliability measures

Reliability is a concept that has different definitions in different contexts. Thus a clear definition of reliability is presented here. An earlier study on the design and validation of advanced driver assistance systems [21] defines reliability in a way originally mentioned [22] as the “probability of a component, subsystem, or complete system, functioning correctly over a given period of time under a given set of operating conditions”. The study by Gietelink [21] also mentions indicators for reliability applicable to warning systems and focuses on the system reliability from the human point of view: the performance of the system visible to the user in terms of true positives (correct activation of the system when needed), true negatives (correct suppression of the safety device), false positives (false alarms due to untimely or incorrect decision of the system) and false negatives (for example, late detections and missed alarms) (Table 2).

Table 2. Prediction matrix with number of samples, categorizations as true negatives, true positives, false negatives or false positives [21], adapted originally from a paper by Lee and Peng [23].

		Actual data	
		Negative (safe)	Positive (threat)
Prediction	Negative (safe)	N_{TN}	N_{FN}
	Positive (threat)	N_{FP}	N_{TP}

These four variables can be used to calculate various reliability measures (Table 3).

Table 3. Reliability measures calculated from the prediction matrix [21].

Rate	**Definition**
Real occurrence rate p	$(N_{FN}+N_{TP})/(N_{TN}+N_{FP}+N_{FN}+N_{TP})$
Accuracy $p_{accuracy}$	$(N_{TN}+N_{TP})/(N_{TN}+N_{FP}+N_{FN}+N_{TP})$
Precision p_{CP}	$N_{TP}/(N_{FP}+N_{TP})$
True positive rate p_{TP}	$N_{TP}/(N_{FN}+N_{TP})$
False negative rate p_{FN}	$N_{FN}/(N_{FN}+N_{TP})$
True negative rate p_{TN}	$N_{TN}/(N_{TN}+N_{FP})$
False positive rate p_{FP}	$N_{FP}/(N_{TN}+N_{FP})$
Reliability p_{rel}	$\sqrt{N_{TP}^2/((N_{FP}+N_{TP})(N_{FN}+N_{TP}))}$

When reliability is defined as above [21], one has to define the operating conditions in which the measurements are made, the length of the observation period, the way the probability of the system functioning correctly is measured, and the criteria for correct functioning of the system.

One of the possible methods for estimating the probability of any event is the relative frequency approach [24]. Basically, this means repeating an experiment many times and calculating the probability of the event by dividing the number of times the event occurred by the number of times the experiment was run. The estimate gains accuracy the more the experiment is run.

The framework described above can be used in situations where four types of outcomes are possible: true positive, false positive, true negative and false negative. In other words, the experiment has a sample space S consisting of TP, FP, TN and FN:

$$S = \{TP, FP, TN, FN\}.$$

In some cases, data collection methods used and characteristics of the system to be evaluated do not allow detection of 'true negative' as a separate event. This may occur in cases where the system under observation and analysis is normally in an idle state producing no output, and observation of the system is performed continuously rather than in event-oriented manner. In these cases the sample space of the experiment consists of three possible outcomes:

$$S = \{TP, FP, FN\}.$$

The prediction matrix must be written without N_{TN} (Table 5). Some reliability measures presented in Table 3 cannot be calculated when no information is available about the number of TN outcomes. Thus real occurrence rate, accuracy, true negative rate and false positive rate were excluded from the group of reliability measures to be calculated (Table 6).

The use of probabilistic techniques such as those described in Tables 2–3 is problematic with events that occur relatively rarely. For example, some potentially hazardous situations may occur only in certain conditions, and for some failures of the system the mean time between failures (MTBF) is longer than or equal to the observation period. However, even if the results are not statistically significant, they can provide useful information about the system.

Reliability block diagrams

Reliability block diagrams [25] have long been used to analyse the reliability of large and complex systems. A reliability block diagram presents a system as a group of components that may be functionally parallel to or in series with each other. Reliability block diagrams can be used together with probabilistic techniques or as tools to understand the behaviour and topology of the system under analysis.

Method chosen

The reliability of the system was defined as the reliability visible to the end-user according to the definition used by Gietelink [21]. Reliability block diagrams were drawn to illustrate the service chain and factors contributing to the reliability of the system. Because no directly applicable definition for a successful use case or acceptable service quality was available in earlier research, a literature study was carried out to establish criteria for a successful use case and to classify the other possible outcomes of a situation in which a train passes a level crossing monitored in the field test. The framework defined by Gietelink [21] had to be adapted before it could be used in the study.

4. MODELLING OF THE SYSTEM

Modelling reliability

Two reliability block diagrams (Figure 3) of the system were drawn based on the technical architecture (Figure 2). Two diagrams instead of one were needed because the service chain is different in some respects for A1 Trax devices and other types of train equipment (CL341 and MC40). The diagram illustrates the components affecting the reliability of the system.

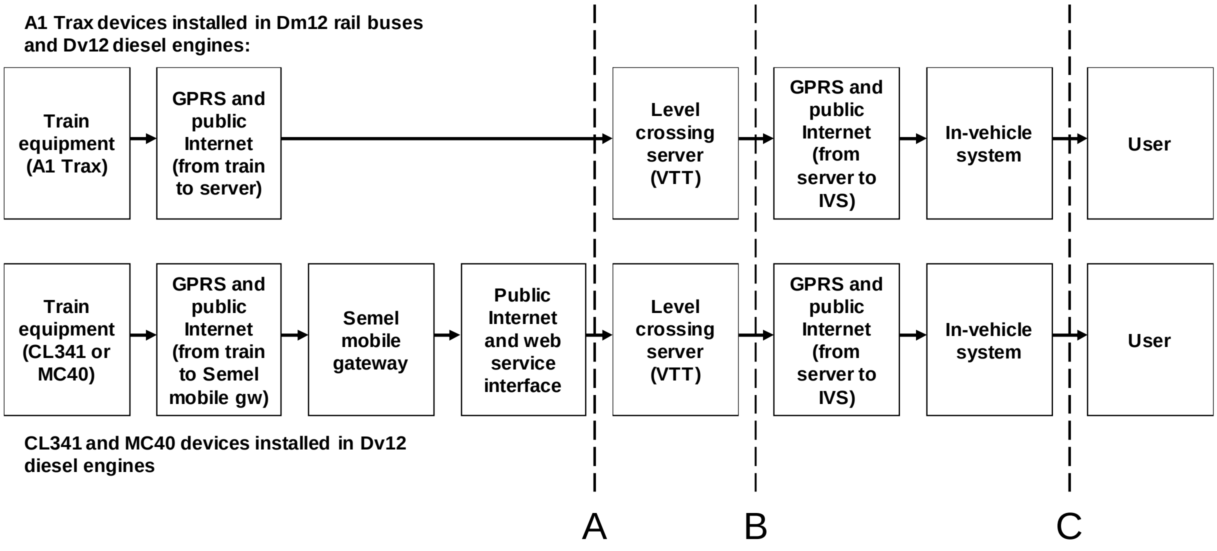

Figure 3. Reliability block diagrams of an in-vehicle warning system for railway level crossings.

Data collection for the reliability analysis was carried out at three points in the service chain: (A) where train position information is received by the level crossing server directly from either the train equipment or the mobile gateway, (B) at the interface providing level crossing status information to the in-vehicle device, and (C) between the in-vehicle system and the end user. Data collection at points A and B was carried out automatically using the data logging features built into the level crossing server software. Data collection at point C was carried out by monitoring the level crossing and the in-vehicle system with video cameras and using a multiplexer to combine the two video streams into the same video data file. Because the focus was on studying reliability from the user point of view, the reliability analysis carried out in this study was performed at point C.

The definition of a successful alarm was created on the basis of the high-level user requirements presented earlier, guidelines available for existing level crossing warning systems and a brief literature study. The correct alarm belonging to category N_{TP} and various types of unsuccessful alarms are illustrated in Figure 4. The other alarms in Figure 4 are classified into category N_{FN} . Even though the system has in fact provided a warning in these cases, classifying these situations as "false negatives" can be reasoned as follows: a warning coming too late, ending too early or being interrupted is obviously outside the stated quality boundaries and may be useless to the driver or cause confusion.

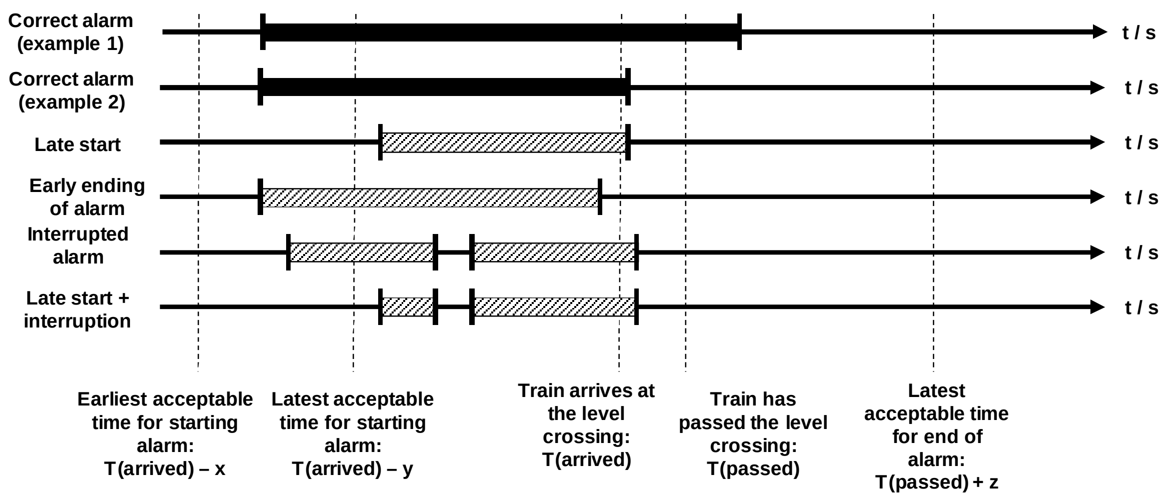

Figure 4. Classification of alarm situations and related errors.

Timing parameters for successful alarm and the discussion related to parameter values are presented in Table 4. The latest acceptable time for starting the alarm must give the driver enough time to react and stop the vehicle. The earliest acceptable time for starting the alarm must be limited to ensure that the warnings are useful to the driver and relevant to his/her driving task. The latest acceptable time must be set in such a way that alarms that are no longer relevant do not confuse users, reduce their trust in the system or cause unacceptable inconvenience.

Table 4. Timing parameters for successful alarm and discussion related to parameter values in Figure 4.

Timing parameters for acceptable alarm		
Parameter	Acceptable values	Estimation of values
x	x ≤ 4 min	Selected on the basis of the usage context of the application, requirements set for roadside level crossing warning systems and a brief literature study. At present, Finnish guidelines for level crossing warning systems do not directly limit the time that level crossing equipment may be in a state of alert before arrival of a train. While a maximum time (40 s) has been determined for trains travelling at the maximum local speed limit [26], these values do not provide guidance for situations in which the train is moving slowly. In other words, a train moving slowly (20 km/h) may keep a level crossing on a 140 km/h railway line in an alert state for several minutes (4 min 40 s). Even longer warning times of up to 5 min have been reported for level crossing warning systems in use [27]. The impact of warning times on driver compliance and means to increase the credibility of warnings provided by level crossing systems have been discussed in a literature review by Yeh and Multer [28]. According to a field study [29] mentioned by Yeh and Multer [28], most violations occurred when the warning time was more than 50 seconds. At level crossings equipped with flashing lights but not gates, violations increased when the warning time exceeded 35 seconds. It can also be argued that a shorter warning time would better reflect the requirements of end users. However, it can also be argued that the system will increase the awareness and alertness of the driver and provide useful information to him/her even in cases where the driver decides not to wait for the train to pass first, and that the driver is not necessarily expected to wait at the level crossing every time he/she receives a warning. When delivering similar information in the same context via different means, it was considered appropriate to use similar or more stringent requirements unless relevant standards or guidelines are available or there are other reasons to choose some other value. On the other hand, it was considered inappropriate to set completely different requirements for the same information delivered with different systems, because this would make the comparison of reliability between existing and novel systems unfair. Less than a minute was considered a target value for maximum warning time, but 4 minutes was selected for this study in order to maintain consistency with requirements set for existing solutions.
y	y ≥ 25 s	According to Finnish guidelines for level crossing warning systems, the minimum time from the start of an alarm to arrival of the train at the level crossing is 20 s [26]. The parameter value in this study was 25 s, as longer times have been used in other countries [2].
z	z ≤ 20 s	The maximum acceptable time from the moment the train has passed the level crossing to the moment the system moves out of the alarm state was estimated on the basis of existing guidelines for level crossing equipment and functionality of the system. For certain types of level crossings, guidelines published by the Finnish Transport Agency identify 5 s as longest acceptable time the level crossing can stay in an alarm state after the train has passed [26]. This requirement applies to both level crossings with boom gates and level crossings with warning lights only. This requirement was considered too stringent because the system under analysis

		provides only a warning but does not prevent the road user crossing the tracks. At present, the warnings given by such a system do not constitute a legal obligation to behave or not in a certain way. In other words, the user of the system is free to make his/her own decision as to whether crossing the tracks is safe after the train has passed, and the only motivation to limit z is to maintain the credibility of the warnings issued by the system. 20 s was considered short enough to effectively limit the number of erroneous alarms given by the system and the inconvenience caused by them to the user.

Data collection

Information about warnings provided by the system and actual trains passing the level crossing was collected with video monitoring. Video streams from two cameras aimed at the display of the in-vehicle system and the level crossing were combined into the same picture and recorded on a hard disk together with a timestamp generated by the hard disk recorder. The collected video data was coded in mpg format with a 25 Hz frame rate and viewed with Avidemux open source video processing software. Various types of events such as all trains that passed the level crossing and all alarms given by the in-vehicle system were documented as a spreadsheet. In total, 500 gigabytes of video data were collected between 7th and 15th March. Data collection was interrupted only when the battery of the monitoring system was changed every day to a fully charged one and when the monitoring system was moved from the Lappohjan satama [Lappohja harbour] to the Skogbyn seisake [Skogby halt] level crossings. These two level crossings are located on a railway line between Hanko and Hyvinkää in Southern Finland (for maps and description in Finnish, see [30]).

The system used for data collection is shown in Figures 5 and 6. Video monitoring was carried out at two level crossings: Lappohjan satama (from 7th to 15^{h} March) and Skogbyn seisake (from 15th to 18th March). The monitoring system with the in-vehicle device was installed on an electric pole about 50-100 metres from the level crossing at both crossings. This distance (50-100 metres) was within the lower distance threshold value mentioned in the system description.

Video monitoring was carried out at two level crossings to reduce the possibility that any unexpected random errors or differences in physical environment would have a disproportionate impact on the results. The observation periods at the two level crossings were different because there was only one set of monitoring and data collection equipment available. When planning the data collection, video monitoring was preferred as a method of data collection for three reasons: First, video recordings were assumed to support the analysis of unsuccessful alarms that may be caused, for example, by certain types of engines or work machines not equipped with the train device. The second reason was the ease of installation and availability of suitable equipment. Third, video recording was considered to be a relatively robust data collection method.

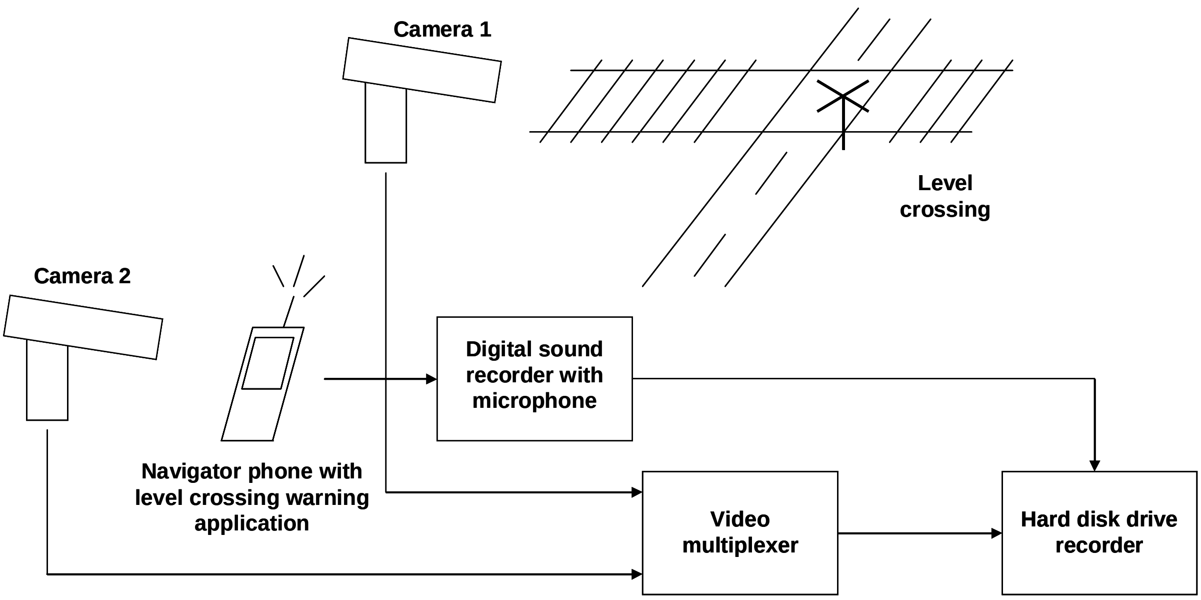

Figure 5. Overview of the system used for video data collection.

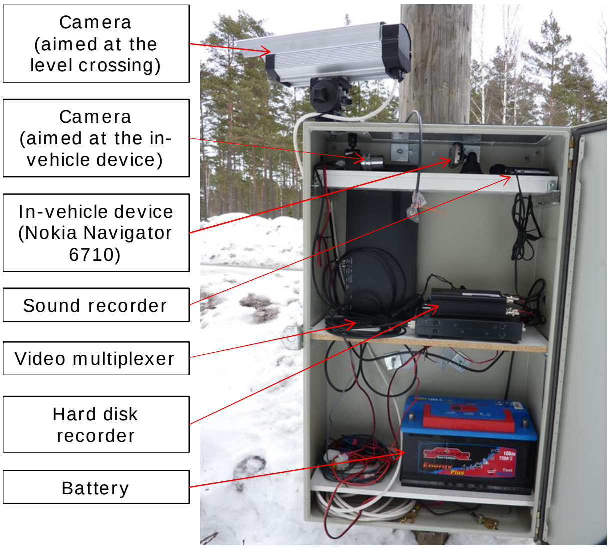

Figure 6. Video monitoring equipment used for video data collection.

5. RESULTS

Reliability results

The video monitoring results are summarised in Table 5.

Table 5. Results of video monitoring at the field test site.

	Passed trains										No train
	No quality deviations	Quality deviations								Missed alarm	False alarm
		Late start	Inter-rupted	Late start + interruption(s)	Late end	Early start	Ended early	Late end + interruptions	Late start + late end		
7.3.2011										7 *	
8.3.2011 00:00-07:46										6 *	
8.3.2011 07:46-13:38	1		1	1						8	
8.3.2011 13:38-24:00										7 *	
9.3.2011										31 *	
10.3.2011										29 *	
11.3.2011 00:00-12:17										14 *	
11.3.2011 12:17-24:00	6	1			1					5	
12.3.2011	6				2			1		10	
13.3.2011	6	2	1		4					5	
14.3.2011	8	2	6		4	1	2		1	6	7
15.3.2011	11	2	4		1		3			7	6
16.3.2011	18		2							2	
17.3.2011	5		2		1					10	
18.3.2011	2		2							3	
* Observations excluded from the results, level crossing server inoperative during the period											

The results of video monitoring (Figures 5-6) have been added to the prediction matrix below (Table 6). A large number of missed alarms occurred early in the test when the service was not operational because of a server failure. The server failure occurred after the system had been verified to be in an operative state by checking that the in-vehicle system received responses to queries it sent to the level crossing server. Failures of the level crossing server are relatively rare events that do not usually occur within 11 days. For this reason, the measured data does not accurately reflect the impact of that failure mode on the reliability of the system. The time during which the server was not operating (6th March at 22:36 – 8th March at 7:46 and 8th March at 13:38 – 11th March at 12:17) was identified on the basis of server log files and video recordings of the display of the in-vehicle client and excluded from the observation period. Any observations recorded during that period were not included when the figures in Table 6 were calculated. The reliability measures calculated from the figures in the prediction matrix (Table 6) are presented in Table 7.

Table 6. Prediction matrix with number of observations categorized as true negatives, true positives, false negatives or false positives, derived from observations in Table 5.

		Actual data	
		No train	Train observed
Prediction	Negative (safe)	Not applicable	103
	Positive (threat)	13	63

Table 7. Reliability measures calculated from the figures in Table 6.

Rate	**Definition**	**Result**
Precision p_{CP}	$N_{TP} / (N_{FP} + N_{TP})$	82.9%
True positive rate p_{TP}	$N_{TP} / (N_{FN} + N_{TP})$	38.0%
False negative rate p_{FN}	$N_{FN} / (N_{FN} + N_{TP})$	62.0%
Reliability p_{rel}	$\sqrt{N_{TP}^2 / ((N_{FP} + N_{TP})(N_{FN} + N_{TP}))}$	56.1%

The share of false positives of all alarms generated by the system is 17.1%.

Confidence intervals

Confidence intervals for the probability of a successful alarm (with 0.95 level of confidence) can be calculated for a probability of p and a sample of n trials with the formula [24]:

$$p_{ci} = p \pm 1.96 \times \sqrt{\frac{p \times (1-p)}{n}} \quad (1), \text{ when } n \times p > 5 \text{ and } n \times (1-p) > 5$$

When using (1) and the figures from Table 5 (Lappohja Harbour from 7th to 15h March and Skogby Halt from 15th to 18th March), it was assumed that each train observed at the same level crossing was equally likely to be successfully detected by the system. The numbers of trains observed, numbers of successful alarms and related values of p and its approximated confidence intervals are listed in Table 8. For both level crossings, the measured success rate and its upper confidence interval are less than the target value.

Table 8. Approximated confidence intervals for true positive rate for both level crossings.

	Skogby Halt	Lappohja Harbour
Trains observed $(N_{TP} + N_{FN})$	36+39=75	27+64=91
Successful alarms (N_{TP})	36	27
p	0.48	0.297
n	75	91
p_{ci}	48%±5.7%	29.7%±4.9%

The upper confidence interval for the number of false alarms can be calculated based on the assumption that they occur randomly and their number within some defined period of time follows the Poisson distribution. The upper confidence interval for the number of false alarms that will occur during the observation period used in the study can then be calculated with the formula:

$a = \frac{1}{2} \times \chi^2_{1-p}$, where $f = 2(x+1)$.

where f is the number of degrees of freedom, p (0.95) is the level of confidence, x (13) is the number of observed events in the sample (observation period), and χ is obtained from a table of upper-tail critical values of chi-square distribution [31]. By using available statistical tables

[32] with the values mentioned above, the number of false alarms during the observation period used is 41 or less, corresponding to a false alarm rate of 41/(63+41)=39.4% with 0.95 level of confidence.

Causes of unsuccessful alarms

Once the results were available, preliminary analysis of the causes of unsuccessful alarms was performed, focusing on situations in which the system had provided no detection at all (56 events, Table 5). Twelve missed alarms were found to be caused by an engine, rail bus or work machine passing the level crossing without a train unit connected to the system. An additional 15 missed alarms were caused by GSM handovers, or by other situations in which the data connection between the train unit and the level crossing server or mobile gateway was disconnected for a short time. Handovers and other brief disconnections in the data connection between the train unit and level crossing server or the mobile gateway were detected by looking at the delay with which the packets were received at the level crossing server (interface A, Figure 3). Log files of the utilization of rolling stock such as rail buses and diesel engines were obtained from the railway operator and matched to observed trains and lists of diesel engines and rail buses equipped with a train unit connected to the system. However, an improved matching procedure is needed to get a more accurate estimate of the number of situations in which the cause of a missed alarm was an unequipped train.

Other identified causes of completely missed alarms were failures of some train units and temporarily occurring situations in which the in-vehicle system was inoperative. It was also suspected that some unsuccessful alarms were related to excessive delay to packets caused by the mobile gateway and the web service interface between the mobile gateway and the level crossing server.

6. DISCUSSION AND CONCLUSIONS

Concluding remarks

The true positive rate was 38.0% if only alarms meeting the previously defined criteria (63 alarms) are classified as successful alarms (Table 7). If also alarms with quality deviations (an additional 47 alarms, Table 5) are counted as true positives, the true positive rate improves to 66.3%. The true positive rate achieved (38%) is considerably smaller than that required from the system (90%). Measurements in the study included data on 166 events over 9 days (Table 5). Even without sophisticated statistical methods, it is possible to conclude that the true positive rate achieved by the system does not meet the required level. Estimation of confidence intervals supports this conclusion.

Many observed situations classified as "false negatives" are actually realised detections with some deviations from the defined quality criteria set, and hence not "totally" missed, and some of them probably still provide useful information to the driver. While some of them may be hazardous to the user in certain situations, others are only a nuisance, such as alarms not ending quickly once the train has passed.

The system generated only 13 false alarms during the observation period of 9 days, which is a relatively small amount when compared to 63 successful alarms and 166 trains observed in total (Table 6). The false alarm rate (39.4%) corresponding to the upper confidence interval for the number of false alarms is less than the value acceptable for the system (60%). Although the results suggest that the false alarm rate achieved would be within the limits acceptable to the system, no final conclusion can be drawn because of the limited empirical data available on the impacts of false alarm rate on user behaviour. Further research is needed to determine a more justified and accurate estimate for the maximum acceptable false alarm rate for the system. Finding a balance between a low false alarm rate and a high probability of detection is a common objective but also a challenge in the design of alarm systems.

The test shows that the expected functionality has been realised, but the true positive rate of the system must be improved. The challenge in further development of the system seems to be the share of false negatives (trains that passed the level crossing but for which the system issued no correct warning) rather than false positive detections (alarms without a train approaching the level crossing).

Video monitoring is a labour-intensive data analysis method. For this reason, only a relatively short observation period of 11 days was feasible. Thus, the results have to be interpreted with caution for all events for which the observation period is shorter than or of the same length as the mean time between a particular type of failure (MTBF) or the mean time to repair it (MTTR). For this reason, the impact of server failures that occurred early in the observation period also had to be excluded from the data to give more accurate information on the reliability of the system in typical operating conditions, and the results obtained do not reflect the impact of the reliability of the level crossing server on the reliability of the system experienced by the end user. The system analysed in this study is a large and complex one that relies on other systems such as mobile and fixed-line communication networks and satellite positioning. This means that there may be failure modes that did not appear within a short observation period at two locations. However, the results most probably reflect the impact of failure modes that occur most frequently and can be expected to have the greatest impact on the reliability of the system. The limitations of the GPRS data connection and the mobile platform used as an in-vehicle device and failed train devices were found to be significant sources of error and randomness in the system.

One can also question whether the criteria for a successful alarm are too broad. For example, warning times of up to 4 minutes may be considered long. The reasons for choosing this parameter value are explained in Table 4. However, choosing tighter criteria for a successful alarm would not have a major impact on the main conclusions, since the true positive rate of the system was already found to have substantial room for improvement with the current parameters. Naturally, the values of parameters x, y and z should be adjusted as more information on user needs and driver behaviour becomes available.

Recommendations for further research

More detailed analysis of the causes of false negatives and false positives would be required to better understand the behaviour of the system and to estimate the potential for improvement. The analysis should make a clear distinction between faults likely to occur during normal operation of the system and causes related to the organisation of the field test such as trains not being equipped with a train unit. It is also recommended that data collection and analysis of reliability be continued once the system has been further developed and improved.

While relatively few problems could be attributed to the accuracy or performance of satellite positioning, the GSM handovers and long round trip times related to GPRS seemed to have a significant impact on the reliability of the system. Therefore, two important questions related to data communication and reliability of the system can be formulated.

The first is whether the limitations of existing GPRS can be overcome with improvements to communication protocols and data-processing algorithms in train units, the level crossing server and the in-vehicle system, and how much they will improve reliability. For example, GSM handovers are likely to occur in the same places, and several types of information can be used to predict the movements of the train during a handover procedure, causing delay or packet loss for the messages sent by the train unit. Different techniques such as matching the observed train to timetables and historical speed profiles of similar locomotives or rail buses in time and space on the same railway line etc. should be tested. In cases where two separate GSM networks are available, two train devices with SIM cards from different network operators could be used in the same engine or rail bus to achieve some redundancy for the GPRS data connection. Using a transmission protocol other than TCP would also shorten the time needed to recover from a disconnection of the data link between the train unit and the level crossing server or mobile gateway.

The second important question is what will be the impact of new data communication services offered by 3G and 4G networks on the reliability of the system. It is probable that shorter data latency and reduced packet loss will improve the operation of the system, but the impact on the system via handover performance is not fully predictable. Empirical tests and detailed analysis of handovers in UMTS and LTE (Long-Term Evolution) and vertical handovers from UMTS or LTE to GPRS offered by 2G networks and vice versa will be needed to get an accurate picture of their impact on the reliability of the system.

It can also be asked whether the reliability targets set for the system in the study are appropriate or optimal in terms of safety impacts. The targets set for the system have been determined on the basis of a literature study and guidelines available for other driving assistance systems. This means that they reflect the current knowledge on the topic but they should not be considered final. The optimum reliability of the system in terms of safety impacts and cost-effectiveness would be an interesting research topic.

The criteria for a successful alarm have been set on the basis of a literature review and existing guidelines for level crossing warning systems. A more detailed analysis should be carried out to continuously improve the values of parameters x, y and z to maximise the safety impacts of the system and to ensure that the system responds to the needs of end users. The study has focused on the reliability of the system but not analysed impacts on user behaviour. Empirical studies of impacts on user behaviour would most likely provide more detailed information on user needs, optimal parameter values and safety benefits of the system.

7. ACKNOWLEDGEMENTS

I wish to thank all my colleagues at VTT who have worked on the project and all the stakeholders who made it possible. Especially, I would like to thank Ms. Marita Hietikko for her valuable comments and Mr. Antti Seise for implementation of the video monitoring system.

8. REFERENCES

[1] European Railway Agency: 'Railway Safety Performance in the European Union 2012' (European Railway Agency, Valenciennes, France, 2012), http://www.era.europa.eu/Document-Register/Documents/SafetyReport2012.pdf, accessed 2nd August 2012

[2] Lazarevic, N., Khoudour, L., El Koursi, E. M., Machy, C., Roberts, C., Impastato, S., Baldassarra, A., Hantak, H., Berrado, A., Cherkaoui, A., Collart Dutileul, S., Hartwig, K., Pelz, M. and Slovak, R.: 'SELCAT D2 - Report about Examination of actual and potential Technologies for Level Crossing', http://www.iva.ing.tu-bs.de/levelcrossing/selcat/lcDocuments/878-26_SELCAT-D2.pdf, accessed 2nd August 2012

[3] Safety Investigation Authority, Finland: 'Teematutkinta tasoristeysonnettomuuksista' [Safety study on level crossing accidents] (Safety Investigation Authority, Helsinki, Finland, 2011), http://www.turvallisuustutkinta.fi/1290610147901 , accessed 6th August 2012

[4]
http://portal.liikennevirasto.fi/sivu/www/f/liikennevirasto/tilastot/onnettomuustilastot/tasoristeysonnettomuudet, accessed 6th August 2012

[5] Elvik, R., and Vaa, T.: 'The handbook of road safety measures' (Elsevier B.V., 2004, Amsterdam, Netherlands)

[6] Öörni, R. and Virtanen, A.: 'In-vehicle warning system for railway level crossings' Proceedings of the 6th European Congress and Exhibition on Intelligent Transport Systems and Services, 2007, Aalborg, Denmark.

[7] 'Pilot Study of Advisory On-Board Vehicle Warning Systems at Railroad Grade Crossings' [Executive summary]. Illinois Department of Transportation, Office of Planning and Programming. Schuamburg, Illinoins, USA; 1997.

[8] Hartong, M., Goel, R., Farkas, C. and Wijesekera, D.: 'PTC-VANET Interactions to Prevent Highway Rail Intersection Crossing Accidents'. Proceedings of the 65th IEEE Vehicular Technology Conference, Dublin, Ireland, 22-25 April 2007.

[9] 'Saving lives at rail crossings' [podcast on the Internet]. Melbourne: La Trobe University; 21st July 2010. Available from: http://www.latrobe.edu.au/news/videos/saving-lives-at-rail-crossings.

[10] Australia. State Government Victoria: 'Australian first: new technology to improve railway crossing safety' [Media release]. Melbourne, Australia, 1st April 2011. Available from: http://www.premier.vic.gov.au/media-centre/media-releases/521-australian-first-new-technology-to-improve-railway-crossing-safety.html [accessed 27th October 2011].

[11] Finnish Transport Agency: 'Varoituslaitosten tekniset toimitusehdot' [Technical requirements for level crossing warning systems] (Finnish Transport Agency, Helsinki, Finland, 2012), http://www2.liikennevirasto.fi/julkaisut/pdf3/ohje_2012_varoituslaitosten_tekniset_web.pdf, accessed 2nd August 2012

[12] Cooper, D. L. and Ragland, D. R.: 'Driver behavior at rail crossings: cost-effective Improvements to increase driver safety at public at-grade rail-highway crossings in California' (California PATH research report UCB-ITS-PRR-2009-24, University of California, Berkeley, USA, 2009),

http://www.path.berkeley.edu/PATH/Publications/PDF/PRR/2009/PRR-2009-24.pdf (accessed 3rd August 2012)

[13] Wullems, C.: ‘Towards the adoption of low-cost rail level crossing warning devices in regional areas of Australia: A review of current technologies and reliability issues’, Safety Science, Vol. 8-9, pp. 1059-1073

[14] Aycin, M. F. and Benekohal, R. F.: ‘Performance Evaluation of the Pilot Study of Advisory On-board Vehicle Warning Systems at Railroad Grade Crossings’ (University of Illinois, Urbana-Champaign, USA, 2002), http://ict.illinois.edu/groups/tol/reports/TOLseries4.pdf, accessed 7th August 2012

[15] Öörni, R., Hietikko, M., Kauvo, K., Lattunen, A. and Virtanen, A.: ’Autossa toimiva junavaroitusjärjestelmä, Kokeilu Hanko–Hyvinkää rataosalla’ [In-vehicle warning system for railway level crossings, pilots tests on Hanko–Hyvinkää railway line] (VTT Technical Research Centre of Finland, Espoo, Finland, 2011), http://www.vtt.fi/inf/pdf/tiedotteet/2011/T2603.pdf (accessed 3rd August 2012)

[16] UNECE, ITS Informal Group: ‘Guidelines on establishing requirements for high-priority warning signals’ (Informal document No. GRPE-61-12, UNECE, ITS Informal Group, 2009), http://www.unece.org/fileadmin/DAM/trans/doc/2011/wp29grpe/ECE-TRANS-WP29-GRPE-61-inf12e.pdf (accessed 3rd August 2012)

[17] Maltz, M. and Shinar, D.: ‘Imperfect in-vehicle collision avoidance warning systems can aid drivers’, Human Factors, Vol. 46, pp. 357-366, 2004

[18] Dingus, T. A, McGehee, D. V., Mankkal, N., Jahns, S. K, Camey, C., and Hankey, J. M.: ‘Human factors field evaluation of automotive headway maintenance/collision warning devices’, Human Factors, Vol. 39, pp. 216-229, 1997

[19] LeBlanc, D., Bezzina, D., Tiernan, T., Freeman, K., Gabel, M., and Pomerleau, D.: ‘System Performance Guidelines for a Prototype Integrated Vehicle-Based Safety System (IVBSS) – Light Vehicle Platform’ (U.S. Department of Transportation, USA, 2008), http://deepblue.lib.umich.edu/bitstream/2027.42/58195/1/100880.pdf (accessed 3rd August 2012)

[20] Gil, M., Multer, J. and Yeh, M.: ‘Effects of active warning reliability on motorist compliance at highway-railroad grade crossings. (U.S Department of Transportation, National

Technical Information Service, Springfield, VA, USA, 2007), http://www.fra.dot.gov/downloads/Research/ord0906.pdf (accessed 3rd August 2012)

[21] Gietelink, O.: 'Design and Validation of Advanced Driver Assistance Systems'. PhD thesis, Delft University of Technology, 2007

[22] Storey, N.: 'Safety-Critical Computer Systems' (Addison-Wesley Longman Ltd, Essex, UK, 1996. ISBN: 0-201-42787-7)

[23] Lee, K. and Peng, H.: 'Evaluation of automotive forward collision warning and collisionavoidance algorithms', International Journal of Vehicle System Dynamics, 2005, 43, (10), pp. 735–751

[24] Milton, J.S., Arnold, J.C.: 'Introduction to Probability and Statistics. Principles and Applications for Engineering and Computing Sciences' (McGraw-Hill, Singapore, 1995)

[25] U.S. Department of Defence: 'Military handbook: Electronic reliability design handbook. Military handbook MIL-HDBK-338B.' (U.S. Department of Defence, USA, Electronic publication, http://webbooks.net/freestuff/mil-hdbk-338b.pdf, accessed 24th February)

[26] Finnish Transport Agency: 'Ratatekniset ohjeet (RATO) osa 6, Turvalaitteet' [Technical guidelines, part 6, Safety systems] (Finnish Transport Agency, Helsinki, Finland, 2012), http://www2.liikennevirasto.fi/julkaisut/pdf3/lo_2012-06_rato6_web.pdf, accessed 2nd August 2012

[27] Lee, J., Nam, D. and Park, D.: 'Analyzing the relationship between grade crossing elements and accidents', Journal of the Eastern Asia Society for Transportation Studies, 2006, Vol. 6, pp. 3658 – 3668, http://www.easts.info/on-line/journal_06/3658.pdf, accessed 3rd August 2012

[28] Yeh, M. and Multer, J.: 'Driver Behavior at Highway-Railroad Grade Crossings: A Literature Review from 1990-2006' (U.S. Department of Transportation, Research and Innovative Technology Administration, John A. Volpe National Transportation Systems Center Cambridge, MA, USA, 2008), http://www.fra.dot.gov/downloads/research/ord0803.pdf [accessed 3rd August 2012]

[29] Bowman, B.L.: 'Analysis of railroad-highway crossing active advance warning devices' Transportation Research Record: Journal of the Transportation Research Board, No. 1114, pp. 141–151, 1989

[30] http://www.tasoristeys.fi, accessed 7th August 2012

[31] Hald, A.: ‘Statistical theory with engineering applications’ (Toppan Company Ltd, Tokyo, Japan, 1952)

[32] http://www.itl.nist.gov/div898/handbook/eda/section3/eda3674.htm, accessed 8th August 2012